



# ***A Latent Trait Model of Simulated Combat Performance***

*Ross R. Vickers, Jr., Ph.D.*



## ***Naval Health Research Center***

---

*Report No. 10-23*

*The views expressed in this article are those of the authors and do not necessarily reflect the official policy or position of the Department of the Navy, Department of Defense, nor the U.S. Government. Approved for public release; distribution is unlimited.*

*This research was conducted in compliance with all applicable federal regulations governing the protection of human subjects in research.*

*Naval Health Research Center  
140 Sylvester Rd.  
San Diego, California 92106-3521*

Pages: 23  
Words: 3458  
Tables: 3  
Figures: 1  
References: 20  
Contact: Ross R. Vickers Jr.  
Email:ross.vickers@med.navy.mil  
Guarantor: Ross R. Vickers Jr.

## **A Latent Trait Model of Simulated Combat Performance**

Ross R. Vickers, Jr., PhD

Warfighter Performance Department  
Naval Health Research Center  
140 Sylvester Rd.  
San Diego, CA 92106-3521

### **KEYWORDS**

Physical Ability, Task Performance, Latent Trait Model

## **ABSTRACT**

Military personnel perform many physically demanding tasks. Identifying the physical abilities that influence performance will contribute to the design of efficient physical training programs. Causal models were constructed to evaluate aerobic capacity (AC), anaerobic power (AP), and muscle endurance (ME) as potential causes of general performance (GP). Five simulated combat tasks defined GP. AP and AC, but not ME, influenced GP. The AP-AC combination contrasted with general strength (GS)-AC models found in earlier studies. No GS measures were available in this study, so the inclusion of AP in the final model may be a case of omitted variable bias. The models to date have consistently excluded ME as a cause. Further study of the importance of AP could be constructive.

## INTRODUCTION

Military personnel perform a wide variety of physical tasks. Different tasks require different physical abilities. Physical training should develop the abilities that have the greatest impact on task performance. Ability-performance modeling provides a means of identifying the relevant abilities and determining their relative impact.

Ability-performance modeling can be carried out at two levels of analysis. Task-level analyses treat each military task individually. Dimension-level analyses combine individual tasks into a general performance (GP) measure. The latter approach yields a single ability-performance model that applies to a wide range of tasks. The alternative of developing a separate model for each task increases the difficulty of characterizing the ability-performance interface.

Several prior studies have demonstrated the viability of dimensional models. General ability dimensions such as general strength (GS) and aerobic capacity (AC) have predicted GP. The resulting models based on general dimensions have adequately summarized the covariation of physical ability tests with task performance.<sup>1-5</sup>

The appropriate level of analysis remains an open question despite recent findings. Those findings are limited to specific combinations of tests and tasks. Extending the coverage of the task domain might demonstrate that task level modeling is appropriate in at least some instances. Toward this end, this study investigated some simulated combat tasks not covered in previous work.

Recently, Harman, Gutekunst, Frykman, Sharp, Nindl, Alemany, and Mello<sup>6</sup> adopted the task-by-task approach to predict performance on four combat activities, a 400-m run, a series of 5 30-m sprints prone to prone, casualty recovery, and obstacle

course performance. Two aspects of their findings stimulated the present re-examination of their evidence. First, the task performance measures were moderately correlated. The correlations could be evidence that the different tasks shared one or more common causal influences.<sup>7</sup> In previous work, analysis of moderately correlated task performance measures has shown that those indicators could be reduced to a single overall performance index.<sup>4,5</sup> Second, Harman et al.<sup>6</sup> constructed a separate predictive model for each of the four combat tasks. The models were based on forward stepwise regression with vertical jump, horizontal jump, sit-ups, push-ups, and a 3.2-km run as potential predictors. The four predictive models contained 10 parameters relating ability tests to the performance of simulated tasks. A model with fewer parameters would be more parsimonious.<sup>3</sup> Previous modeling efforts suggest that as few as 2 parameters can adequately characterize the ability-performance interface.

The present reanalysis of Harman et al.'s<sup>6</sup> data addressed the major questions arising from the preceding observations. First, can performance be represented as a GP dimension? Second, which ability dimensions affect GP? Finally, does the model based on general dimensions adequately account for the relationships of specific tests with specific tasks?

## **METHODS**

### *Data Source*

The analyses examined the covariance matrix for tests and tasks generated from the standard deviations and correlations reported in Tables 1 through 4 of Harman et al.<sup>6</sup> The statistics summarized test results for a sample of 32 physically-trained men.

### *Measurements*

The physical ability tests included measures of vertical jump performance, horizontal jump performance, sit-ups, push-ups, and a 3.2-km run. The simulated combat tasks included a 400-m run, a series of 5 30-m sprints starting and ending in the prone position on each sprint, casualty evacuation, and an obstacle course. Detailed descriptions of the measurement procedures can be found in Harman et al.<sup>6</sup>

### *Analysis Procedures*

Structural equation models (SEMs) were constructed with the LISREL 8.5 computer program (Scientific Software International, Chicago, IL). The modeling procedure began with separate analyses to construct measurement models for physical ability and performance. The measurement models then were combined to construct a path model with ability measures as causes of performance. This two-step procedure separated the construction of the auxiliary measurement models from substantive hypothesis tests<sup>9</sup>. Following McDonald and Ho<sup>10</sup>, the presentation and discussion of study findings uses the terms “measurement model” and “path model” to differentiate the two types of model. The path models consist of the hypothesized causal effects of physical abilities on simulated combat performance.

A three-dimensional ability model was constructed. The vertical jump and horizontal jump defined one dimension. Sit-ups and push-ups defined the second dimension. The 3.2-km run defined the third dimension. These dimensions corresponded to Anaerobic Power (AP), Muscle Endurance (ME), and Aerobic Capacity (AC) dimensions identified in previous studies.<sup>4,5</sup>

Fixing the variances at 1.000 established the scales for the latent traits representing the general ability and GP dimensions. This method of scaling made it possible to estimate factor loadings for each indicator variable in the measurement models. The alternative approach of fixing one factor loading at 1.00 would have meant that the relevance of at least one indicator to the latent trait was simply assumed. It then would be impossible to test for the appropriateness of assigning the chosen scaling indicator to the trait. A formal test for the relevance of every indicator was preferable. The second scaling decision involved error terms in the measurement models. The parameter estimates for the initial measurement models included some negative error variances. Negative variances are impossible by definition, so the negative error estimates must have been a result of sampling error. Because the true variances must have been greater than zero, substituting zero for the negative values provided an estimate that must have been closer to the true error variance (Table I).

The error variance for the 3.2-km run in the ability measurement model was fixed at zero for a different reason. In this case, there was only one indicator variable to define the hypothesized latent trait. Fixing the error variance at zero meant that the aerobic capacity trait was identical to performance on the 3.2-km run. The strong relationship of performance on distance runs with laboratory measurements of maximal oxygen uptake, the gold standard for cardiorespiratory justified this decision.<sup>11-15</sup> It should be noted, however, that the latent trait defined by the 3.2-km run could also be interpreted simply as distance running performance. Unpublished factor analyses of run tests covering varying distances showing that long runs (i.e., >2 km) defined a single general performance factor provide support for this alternative explanation.

Model evaluation criteria included the model  $\chi^2$ , the Tucker-Lewis index (TLI), which is also known as the non-normed fit index, the standardized root mean square residual (SRMR), and critical N (see Arbuckle & Wothke<sup>16</sup> for definitions). Correspondence with prior research findings was an additional consideration in the final model selection.

## RESULTS

### *Performance Measurement Model*

Harman et al.<sup>6</sup> reported moderate correlations among the performance measures in Table III of their paper. A single dimension adequately summarized the covariation of those measures (see Table I). The residual covariation among the measures was not statistically significant ( $\chi^2 = 1.64$ , 1 *df*,  $p > .440$ ). All standardized residuals were small ( $|z| < 1.28$ ). The unidimensional model satisfied two widely-used goodness-of-fit criteria for structural equation modeling (i.e., TLI  $> .900$  and SRMR  $< .05$ ). However, the model only approached the recommended critical N criterion (i.e.,  $N > 200$ ).

The GP measurement model could have been simplified further. The error variances for the 400-m run and the 30-m rushes could have been fixed at zero. Those error terms were positive, but the  $t$  values did not meet the  $|t| > 2.00$  criterion that is the usual justification for retaining a parameter in a structural model. The empirical error estimates were retained because the variance estimates were positive. A small positive variance was more plausible than a zero variance.

### *Ability Measurement Model*

The ability measurement model was correctly specified (see Table II). All factor loadings were large enough to be retained, i.e.,  $t > 2.00$ . No modification index for the model approached significance, so there was no reason to remove the constraint on any factor loading that had been fixed at zero.

The ability measurement model accounted for the covariation of the ability tests. The model provided a significantly better fit to the data than a null model,  $\Delta\chi^2 = 53.66, 9 df, p < .001$ . The residual covariation was not statistically significant,  $\chi^2 = 3.15, 1 df, p > .075$ , SRMR was  $< .05$ , and TLI was  $> .900$ , and critical N was close to the criterion value of 200.

The correlations between physical ability dimensions were statistically significant. By Cohen's<sup>17</sup> criteria, the relationship between AP and ME,  $r = .532, SE = .215, t = 2.48$ , was moderately large, as was the relationship between AP and AC,  $r = -.420, SE = .111, t = -3.80$ . The very large correlation of ME with AC indicated virtual identity of the two latent traits,  $r = -.966, SE = .109, t = -8.89$ .

#### *Ability-Performance Path Model*

The analyses of ability-performance relationships produced a set of equivalent models (see Table III). Two models are equivalent if they achieve equal explanatory or predictive power with the same number of parameters.<sup>18</sup> Sampling variation makes literal equivalence unlikely in empirical analyses even if the true underlying models are equivalent. For this reason, identifying path models that are approximately equivalent is more useful than limiting focusing on literally equivalent models.

Equivalent model identification proceeded in two steps. The first step grouped models based on the number of causal parameters. The second step compared the explanatory power of alternative models that had equal numbers of causal parameters. Models of equal parametric complexity were considered approximately equivalent if there was little difference in explanatory power. Table 3 presents the findings for the 7 alternative ability-GP models. Three models contained a single causal effect of ability on GP. Three models contained two causal effects, and one model contained 3 causal effects.

The ME model would be favored over the other single effect choices based on a larger reduction in  $\chi^2$  relative to the null model, a smaller SRMR, a larger TLI, a stronger causal effect on GP, and a larger  $R^2$ . However, the differences between the ME model and the AP and AC models were small. The  $\chi^2$  values differed by  $\leq 1.34$  and SRMRs were similar. TLI values differed moderately, but this difference may not be important because the TLIs were computed from virtually identical  $\chi^2$ s. The estimated effect of ability on GP appeared to differ between models,  $b = -.676$  to  $b = -.601$ , but the differences were small relative to the standard errors for those parameters,  $.185 \leq SE \leq .200$ . However, if a single-predictor model had to be selected, the ME model would be preferred because it fared slightly better than the alternatives on every model evaluation criterion.

Adding a second ability-GP effect produced a slight improvement in the overall fit of the model relative to the single predictor models,  $\Delta\chi^2 < 2.07$ . Despite the modest improvements in overall fit, the  $R^2$  for the GP latent trait increased enough to indicate effects that Cohen<sup>17</sup> would classify as small, but potentially important. Thus, two predictor models merited further examination.

The three models with two causal parameters provided comparable accounts of the ability test-performance task covariation. The  $\chi^2$  values were comparable for all three models, and TLI differed only slightly. All SRMR values exceeded .05, but model differences were slight.

Additional criteria favored the AP + AC model. The ME + AC model could be ruled out because the estimated effect of ME was impossibly large and because neither of the estimated effects of ability on GP was statistically significant (i.e.,  $t < 2.00$ ).<sup>3</sup>

The AP + ME model was ruled out for a different reason. Only ME was a significant predictor of GP. Dropping the hypothesized effect of AP because it was not statistically significant would reduce the AP + ME model to the ME model. The model selection problem would revert to choosing among the single predictor models.

The AP + AC model avoided the difficulties of the other two-parameter models. Both abilities produced reasonable effects on GP, so there was justification for a two-parameter model. Also, the  $R^2$  for the AP + AC model was larger than that for the best one-parameter model.

The three-dimensional model was not a competitive alternative. This model did not improve on the goodness of fit of the AP + AC model. All three hypothesized causal effects were statistically nonsignificant. TLI was substantially less than TLI for the two-predictor models. SRMR equaled the SRMR for two-predictor models.

The three-predictor model did produce one noteworthy model choice observation. The estimated effects of AP,  $b = -.397$ , and AC,  $b = .357$ , were very close to the corresponding estimates in the AP + AC model. Both effects were much stronger than the effect estimated for ME,  $b = -.102$ . Explanatory models often are constructed by entering

all possible predictors into an initial model. The initial model then is simplified by eliminating statistically nonsignificant predictors until only significant predictors remain. Applying this practice to the present data would produce the AP + AC model. Thus, the three-parameter path model provided additional justification for adopting the AP + AC model.

Figure 1 presents the major elements of the AP + AC model. The error terms for the model and the correlations among the ability latent traits have been omitted to focus attention on the definitions of the latent traits and the causal effects of ability on GP.

### *Residuals Analysis*

The third research question, “Does the model based on general dimensions adequately account for the relationships of specific tests with specific tasks?” was addressed by analyzing the residual covariances. Large residual covariances would have been found if the latent trait model failed to account for the covariation of specific physical ability tests with specific performance tasks. For example, it might be reasonable to expect that the general model would not fully account for the covariation of the 3.2-km run with the 400-m run. Both the nominal test and the nominal task involve running, so any variation that was specific to running would result in a large residual<sup>9</sup>.

There were no strong residual associations. This conclusion was reached based on the standardized residuals. Given the general assumption that greater physical ability will lead to better performance, meaningful residuals would be positive. In fact, only 11 of the 20 standardized residuals in this study (5 tests x 4 tasks) were positive. None of the standardized residuals was statistically significant;  $z$ -scores ranged from  $z = -1.73$ ,  $p > .083$ , two-tailed, to  $z = 1.80$ ,  $p > .071$ , two-tailed.

Modification indices provide a different perspective on the residuals problem. These indices are estimates of how much the overall fit of the model would be improved if a constrained parameter were freely estimated. Large modification indices would indicate that the constraints results in a misspecified model.

The results were ambiguous with respect to possible model misspecification. A Bonferroni adjustment to the statistical significance criterion of  $p < .0025$  was introduced to allow for the fact that 20 modification indices were considered. Four modification indices would have been statistically significant by the usual  $p < .05$  criterion: 3.2-km run/400-m run ( $\chi^2 = 5.97, p < .015$ ); 3.2-km run/50-yd rush ( $\chi^2 = 4.54, p < .034$ ); push-ups/casualty evacuation ( $\chi^2 = 6.70, p < .010$ ); horizontal jump/obstacle course ( $\chi^2 = 3.89, p < .049$ ). However, no modification index was large enough to satisfy the Bonferroni criterion (Figure I).

Further examination of the modification indices that met the  $p < .05$  criterion raised additional doubts about the appropriateness of adding any model parameters linking specific ability tests to specific tasks. The LISREL program estimates the parameter value that would result if a constrained parameter were freely estimated. In the present case, two of the estimated changes linked greater ability to better performance: 3.2-km run time with 400-m run time,  $r = .12$ ;<sup>5</sup> horizontal jump with obstacle course performance,  $r = -.15$ . The other two parameter estimates paired higher ability with poorer performance: 3.2-km run with 30-m rushes,  $r = -.099$ ; push-ups with casualty evacuation times,  $r = .326$ . When the overall pattern of evidence was considered, the modification indices showed that the residual associations that were small, implausible, or both.

## DISCUSSION

This reanalysis of Harman et al.'s<sup>6</sup> evidence addressed three questions. First, is GP a sound representation of task performance? Second, which physical abilities affect GP? Finally, can general ability and performance constructs adequately account for the relationships of scores on specific ability tests with performance on specific military tasks? The evidence provided a basis for answering each question.

Is GP a sound representation of task performance? The apparent distinctiveness of combat tasks suggests that the answer to this question should be no. However, the moderately strong relationships between tasks defined a single general performance capability. Performance on different tasks defined a single performance dimension and each task was significantly related to that dimension. This result replicated previous findings with different military task sets.<sup>1,2,4,5</sup>

Which physical abilities affect GP? Seven causal models were constructed to answer this question. All of the models had approximately equal explanatory power. Nevertheless, several criteria indicated that the ME and AP + AC models were marginally superior to the other five models. The AP + AC model was the better choice despite its relative lack of parsimony. When AP, AC, and ME were included in the same model, AP and AC effects on GP were moderately large, while the ME effect was just large enough to avoid being classified as trivial. ME was positively correlated with AP and AC, so the explanatory power of the ME model could represent omitted variable bias.<sup>19</sup> Assuming AP and AC were the true causes of differences in GP, the apparent effect of ME on GP was inflated because the estimate incorporated part of the causal effect of AP and part of the causal effect of AC.

Previous work strengthens the argument for the AP + AC model. That work identified general strength (GS) and AC as the causes of general differences in military task performance. ME did not enter into the causal models. Furthermore, the correlation of AP with GS was moderate or large. AP was not related to GP after controlling for its relationship to GS. In this study, the AP + AC model was the closest possible approximation to the GS + AC models. Combining the results of this study with those of earlier studies, the apparent AP effect on GP in this study could represent omitted variable bias.

Study limitations should be noted. The absence of GS measures has been noted. The small sample size was another limitation that reduced the power of the statistical tests. This problem was not important for measurement models, because all of the factor loadings were significant despite the small sample size. However, larger samples would have sharpened the comparison of path models by amplifying the differences in the associated  $\chi^2$  values. Finally, the performance measures were simulated battlefield tasks. It cannot be taken for granted that the results will generalize to actual performance in a combat setting (M. Sharp, personal communication, 14 January 2010).

This treatment of Harman et al.'s<sup>6</sup> model complements their work. Their study was designed to identify field-expedient ability tests that predicted performance. Their study achieved its objective, but extending the treatment of the data to formulate causal models has additional benefits. The extension highlights the need for GS measures to ensure accurate inferences about performance. Future studies should pursue this end by employing a well-defined ability measurement model that covers the full range of physical abilities.<sup>20</sup> Accurate identification of the physical abilities that contribute to

military task performance will reduce the risk of developing misguided physical training programs. The programs can be designed to develop critical abilities and to measure progress using performance-relevant ability tests. The results of this study were consistent with the findings from previous work indicating that general dimensions provide the appropriate level of analysis for modeling the relationship of physical abilities with performance. The implication is that training programs should be designed to promote general capabilities such as AP, GS, and AC.

## **ACKNOWLEDGMENT**

This represents the Naval Health Research Center (NHRC) report 10-23, supported by the Office of Naval Research, under work unit no. 6704. This research has been conducted in compliance with all applicable federal regulations governing the protection of human subjects in research.

## REFERENCES

1. Vickers, R. R., Jr.: Physical task performance: complexity of the ability-performance interface, Technical Report No. 95-30. San Diego, CA: Naval Health Research Center, 1995.
2. Vickers, R. R., Jr.: Generalizability test of a physical ability-job performance model, Technical Report No. 96-16. San Diego, CA: Naval Health Research Center, 1996.
3. Vickers, R. R., Jr.: Physical strength and the performance of moderate duration tasks, Technical Report No. 03-08. San Diego, CA: Naval Health Research Center, 2003.
4. Vickers, R. R., Jr.: Physical Abilities and Military Task performance: A Replication and Extension, Technical Report No. 09-30. San Diego, CA: Naval Health Research Center, 2009.
5. Vickers, R. R., Jr., Hodgdon, J. A., & Beckett, M. B.: Physical Ability-Task Performance Models: Assessing the Risk of Omitted Variable Bias, Technical Report No. 09-04. San Diego, CA: Naval Health Research Center, 2009.
6. Harman EA, Gutekunst DJ, Frykman PN, Sharp, MA, Nindl, BC, Alemany, JA, et al: Prediction of simulated battlefield physical performance from field-expedient tests, *Mil Med* 2008; 173:36-41.
7. Glymour C, Scheines R, Spirtes P, Kelly K: Discovering causal structure: artificial intelligence, philosophy of science, and statistical modeling, pp15-59. Orlando, FL, Academic Press, 1987.
8. Popper KR: The logic of scientific discovery, pp 112-145. New York, Basic Books, 1959.

9. Meehl PE: Appraising and amending theories: the strategy of Lakatosian defense and two principles that warrant it, *Psychological Inquiry* 1990; 1:108-41.
10. McDonald RP, Ho M-HR: Principles and practice in reporting structural equation analyses, *Psychol Methods* 2002; 7:64-82.
11. Knapik J: The Army Physical Fitness Test (APFT): A Review of the Literature, *Mil Med* 1989; 154:326-9
12. Safrit MJ, Hooper LM, Ehlert SA, Costa MG, Patterson P: The validity generalization of distance run tests, *Canadian Journal of Sports Science* 1988; 13:188-96.
13. Vickers, R. R., Jr.; Running performance as an indicator of VO<sub>2</sub>max: a replication of distance effects, Technical Report No. 01-24. San Diego, CA: Naval Health Research Center, 2001.
14. Vickers, R. R., Jr.: Running performance as an indicator of VO<sub>2</sub>max: distance effects, Technical Report No. 01-20. San Diego, CA: Naval Health Research Center, 2001.
15. Vickers, R. R., Jr.: Modeling test run validity: a meta-analytic approach, Technical Report No. 02-27. San Diego, CA: Naval Health Research Center, 2002.
16. Arbuckle JL, Wothke W: Amos 4.0 user's guide, pp395-416. Chicago, SmallWaters Corporation, 1999.
17. Cohen J: Statistical Power Analysis for the Behavioral Sciences. Second ed. Hillsdale, NJ, Erlbaum, 1988.
18. MacCallum RC, Wegener DT, Uchino BN, Fabrigar LR: The problem of equivalent models in applications of covariance structure analysis, *Psychol Bull* 1993; 114:185-99.
19. James LR, Mulaik SA, Brett J: Causal analysis: assumptions, models, and data, pp 71-80. Beverly Hills, CA, Sage Publications, 1982.

20. Myers DC, Gebhardt DL, Crump CE, Fleishman EA: The dimensions of human physical performance: factor analyses of strength, stamina, flexibility, and body composition measures, Human Performance 1993; 6:309-44.

## Tables

Table I.

*GP Measurement Model*

	LY	SE(LY)	<i>t</i> -value	TE	SE(TE)	<i>t</i> -value
400-m run	8.47	1.34	6.34	14.82	7.56	1.96
Repeated sprints	9.86	1.43	6.88	8.78	9.15	.96
Casualty evacuation	1.65	.61	2.72	9.51	2.46	3.87
Obstacle course	8.55	1.97	4.34	78.15	21.18	3.69

Note. LY is the loading of the indicator variable on the latent trait. TE is the residual variance for the indicator.

Table II.

*Ability Measurement Model*

	LX	SE(LX)	<i>t</i> -value	TD	SE(TD)	<i>t</i> -value
<b><i>Anaerobic power</i></b>						
Vertical jump	5.25	1.15	4.58	27.16	6.90	3.94
Horizontal jump	25.60	3.25	7.87	– – <sup>a</sup>	– – <sup>a</sup>	– – <sup>a</sup>
<b><i>Muscle endurance</i></b>						
Push-ups	6.09	2.65	2.30	79.53	31.32	2.54
Sit-ups	7.26	3.01	2.41	84.19	40.17	2.10
<b><i>Aerobic capacity</i></b>						
2-mi run	105.30	11.47	9.18	– – <sup>b</sup>	– – <sup>b</sup>	– – <sup>b</sup>

Note. LX is the loading of the indicator variable on the latent trait. TD is the residual variance for the indicator.

<sup>a</sup>TD was fixed at .000 because the initial analysis indicated that this parameter was negative. This result presumably was a random sampling effect, but negative variances are not meaningful, so the true variance clearly was underestimated. <sup>b</sup>TD was fixed at .000 because there was only a single indicator. Therefore, the latent trait was identical to the indicator variable.

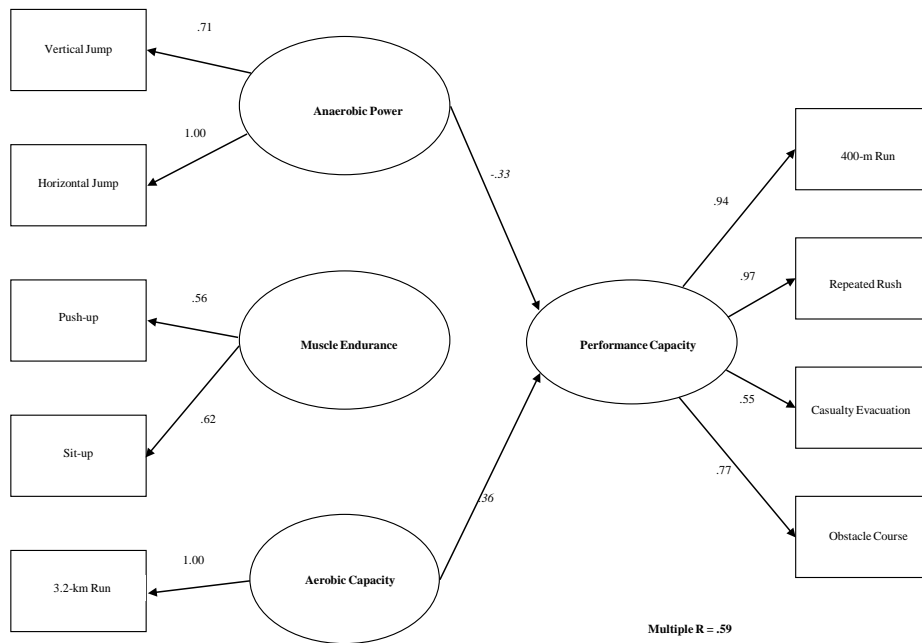
Table III.

*Path Models Relating Ability to GP*

Model	$\chi^2$	SRMR	$\Delta\chi^2$ <sup>a</sup>	df	Sig <sup>b</sup>	TLI	AP	Estimated Causal Effects		R <sup>2</sup>
								ME	AC	
Null	56.69	.325								
AP	50.36	.182	6.33	1	.012	.123	-.601**			.266
ME	49.86	.179	6.83	1	.009	.140		-.676**		.313
AC	51.20	.191	5.49	1	.019	.096			.622**	.279
AP+ME	48.30	.160	8.39	2	.015	.111	-.344	-.483*		.346
AP+AC	48.27	.160	8.42	2	.015	.113	-.413*		.449*	.345
ME+AC	48.90	.174	7.79	2	.020	.091		-1.250	-.585	.329
All	48.28	.160	8.41	3	.038	.023	-.397	-.102	.357	.346

<sup>a</sup>Improvement in model fit relative to the null model. <sup>b</sup>Relative to the null model.\*  $|t| > 2.00$ .\*\*  $|t| > 3.00$ .

Figure I.



Note. Error terms are not shown.

Figure I. Best Ability-GP Model

## REPORT DOCUMENTATION PAGE

The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing the burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB Control number. **PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.**

<b>1. REPORT DATE (DD MM YY)</b> 17 02 10		<b>2. REPORT TYPE</b> Journal Submission		<b>3. DATES COVERED (from – to)</b> Nov 09 – Jan 10	
<b>4. TITLE</b> A Latent Trait Model of Simulated Combat Performance				<b>5a. Contract Number:</b> <b>5b. Grant Number:</b> <b>5c. Program Element Number:</b> <b>5d. Project Number:</b> <b>5e. Task Number:</b> <b>5f. Work Unit Number:</b> 60704	
<b>6. AUTHORS</b> Vickers, Ross R. Jr.				<b>8. PERFORMING ORGANIZATION REPORT NUMBER</b>  Report No. 10-23	
<b>7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)</b> Commanding Officer Naval Health Research Center 140 Sylvester Rd San Diego, CA 92106-3521					
<b>8. SPONSORING/MONITORING AGENCY NAMES(S) AND ADDRESS(ES)</b> Commanding Officer Naval Medical Research Center 503 Robert Grant Ave Silver Spring, MD 20910-7500				<b>10. SPONSOR/MONITOR'S ACRONYM(S)</b> NMRC/NMSC	
Commander Navy Medicine Support Command P.O. Box 140 Jacksonville, FL 32212-0140				<b>11. SPONSOR/MONITOR'S REPORT NUMBER(s)</b>	
<b>12. DISTRIBUTION/AVAILABILITY STATEMENT</b> Approved for public release; distribution is unlimited.					
<b>13. SUPPLEMENTARY NOTES</b>					
<b>14. ABSTRACT</b>  Military personnel perform many physically demanding tasks. Identifying the physical abilities that influence performance will contribute to the design of efficient physical training programs. Causal models were constructed to evaluate aerobic capacity (AC), anaerobic power (AP), and muscle endurance (ME) as potential causes of general performance (GP). Five simulated combat tasks defined GP. AP and AC, but not ME, influenced GP. The AP-AC combination contrasted with general strength (GS)-AC models found in earlier studies. No GS measures were available in this study, so the inclusion of AP in the final model may be a case of omitted variable bias. The models to date have consistently excluded ME as a cause. Further study of the importance of AP could be constructive.					
<b>15. SUBJECT TERMS</b> physical ability, task performance, latent trait model					
<b>16. SECURITY CLASSIFICATION OF:</b>			<b>17. LIMITATION OF ABSTRACT</b>	<b>18. NUMBER OF PAGES</b>	<b>18a. NAME OF RESPONSIBLE PERSON</b>
<b>a. REPORT</b> UNCL	<b>b. ABSTRACT</b> UNCL	<b>c. THIS PAGE</b> UNCL	UNCL	15	Commanding Officer
					<b>18b. TELEPHONE NUMBER (INCLUDING AREA CODE)</b> COMM/DSN: (619) 553-8429